A PRE-OPERATIVE PREDICTIVE MODEL FOR THE CLASSIFICATION OF NEWLY DIAGNOSED RENAL MASSES LESS THAN 5 CM IN DIAMETER AS BENIGN OR MALIGNANT

by

Ricardo A. Rendon

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF COMMUNITY HEALTH AND EPIDEMIOLOGY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled “A PRE-OPERATIVE PREDICTIVE MODEL FOR THE CLASSIFICATION OF NEWLY DIAGNOSED RENAL MASSES LESS THAN 5 CM IN DIAMETER AS BENIGN OR MALIGNANT” by Ricardo A. Rendon in partial fulfilment of the requirements for the degree of Master of Science.

Dated: August 15, 2012

Supervisors: _____

_____

Reader: _____

DALHOUSIE UNIVERSITY

DATE: August 15, 2012

AUTHOR: Ricardo A. Rendon

TITLE: A PRE-OPERATIVE PREDICTIVE MODEL FOR THE CLASSIFICATION OF NEWLY DIAGNOSED RENAL MASSES LESS THAN 5 CM IN DIAMETER AS BENIGN OR MALIGNANT

DEPARTMENT OR SCHOOL: Department of Community Health and Epidemiology

DEGREE: MSc CONVOCATION: October YEAR: 2012

_____

Signature of Author

Table of Contents

List of Tables

List of Figures

Abstract

Objective: To develop a predictive model for preoperative differentiation between benign (B) and malignant (M) histology in patients with renal masses (RM) using recursive partitioning.

Methods: We analyzed preoperative patient and tumour characteristics in 395 subjects who had surgery for RM suspicious for renal cell carcinoma.

Results: The model predicted B vs. M histology with an overall accuracy of 89.6% (95% CI 86.2,92.5). It assigned patients with smaller tumours (<5.67cc) and a predominantly (>45%) exophytic component a high risk of B disease (52.6%). Patients with symptoms, larger tumours (>5.67cc) and larger endophytic component (>35%) have a 0% risk of B disease.

Conclusion: B vs. M disease can be predicted accurately. This predictive accuracy is higher than that shown in renal biopsy series. It is hypothesized that for smaller and exophytic RMs, a biopsy is indicated. Symptomatic, larger and endophytic RMs should be removed without further investigations.

List of Abbreviations Used

AJCC	=	American Joint Committee on Cancer
B	=	Benign
CART	=	Classification and Regression Tree
CI	=	Confidence Interval
cm^3	=	Cubic Centimeters
cm	=	Centimeters
CT	=	computed tomography
Endo	=	Degree of Endophytic Component
M	=	Malignant
MRI	=	Magnetic Resonance Imaging
N	=	Number of Observations
OR	=	Odds Ratio
PMH	=	Princess Margaret Hospital
R cm^3	=	Renal Cell Carcinoma
UICC	=	International Union Against Cancer
USA	=	United States of America
Vol	=	Tumour Volume at Diagnosis
y	=	Years of Age

Acknowledgements

I would like to express my gratitude to my supervisors, Mohamed Abdolell and Susan Kirkland, whose expertise, understanding, and patience, added tremendously to my graduate experience. I appreciate their vast expertise and skills in different areas, which combined added great depth to this project and my knowledge overall.

Chapter One: Introduction

1.1. The Diagnostic and Treatment Dilemma of Small Renal Masses

Renal cell carcinoma (RCC) is the most common malignancy of the kidney. Renal cell carcinoma usually presents incidentally on abdominal imaging performed for various reasons. The incidence of small, incidentally detected renal masses presumed to be renal cell carcinoma is rising.[1] The only proven curative treatment for RCC is surgical removal with partial or radical nephrectomy. These procedures have a significant complication rate and affect the renal function. Traditionally, newly diagnosed renal masses are surgically removed based on image findings only and are infrequently biopsied before surgery. Upon surgical removal of these renal masses, up to 48% of them have proven to be benign, thereby exposing these patients to unnecessary risks. Currently, there are no adequate pre-operative predictors of benign or malignant disease for these radiologically detected renal masses. The only decision tool available for clinical use is a complex nomogram that is rarely utilized and uses variables that have limited predictive ability. The purpose of this study was to develop a classification tree in order to aid in the pre-operative assessment of which patients have an increased risk of harbouring a benign renal mass. It is anticipated that a classification tree with a good overall predictive accuracy will be more user friendly for clinicians and therefore will be utilized more often.

1.2. Hypothesis

Using pre-operative prognostic factors, a predictive model can be developed that can accurately classify renal masses < 5 cm, presumed to be renal cell carcinoma, as benign or malignant.

1.3. Primary Objective

To develop a pre-operative classification tree to determine which renal masses <5 cm presumed to be renal cell carcinoma are in fact malignant or benign.

1.4. Secondary Objectives

1.4.1. To validate the classification tree

1.4.2. To evaluate the predictive ability of the classification tree

Chapter Two: Literature Review

2.1. Descriptive Epidemiology of Kidney Cancer

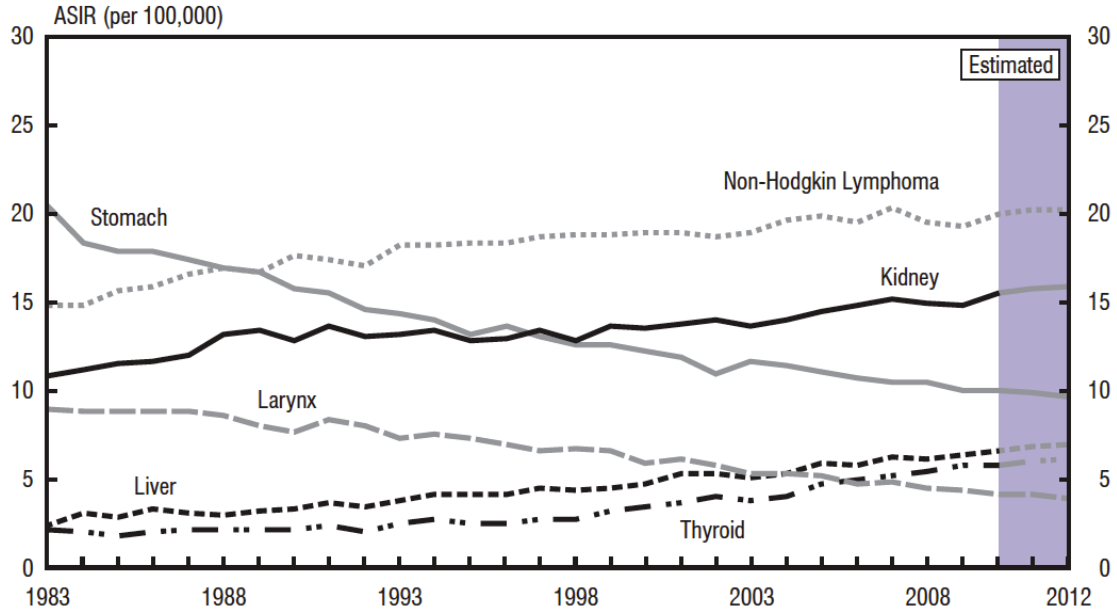
Renal Cell Carcinoma is the most common malignancy of the kidney. In Canada, kidney cancer accounts for approximately 3.5% of all adult male cancers and 2.5% of all adult female new cancer diagnoses.[1] Kidney cancer also accounts for 2.7% and 1.7% of all adult male and female cancer deaths respectively. Close to 5,600 new cases are expected in Canada in 2012, with an estimate of 1,700 resulting deaths. The lifetime probability of developing renal tumours amongst all races in Canada is 1.6% while the lifetime probability of dying from renal tumours is 0.7% with an approximate mortality of 43% at 5 years.

The incidence of RCC in Canada and the United States of America (USA) is increasing (Figure 1.).[1-3] In Canada, the annual percent change in age-standardized incidence rates between 1998 and 2007 for males was 2.6 and 1.9 for females while the annual percent change in age-standardized mortality rates between 1998 and 2007 for males was -0.8 and -0.9 for females. In Canada and the USA, the majority of new RCCs are asymptomatic and are incidentally discovered. The TNM classification developed by the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC) guidelines is the most commonly used system to stage tumour status.[4] Clinical stage

migration to localized stage T1N0M0RCCs has been observed over the past 30 years.[3] Patients are now diagnosed with smaller tumours. Patients in the 7th-9th decades of life have experienced the largest increase in incidence with a ten-fold rise between 1935-39 and 1985-89.[3] There are at least two possible explanations for this increase. First, there is an artefact of increased imaging studies as slow growing lesions, which may have been asymptomatic for years, are now more likely to be detected through X-rays while still small. This hypothesis is based on earlier series. Second, two recent studies have demonstrated a true increase in incidence as evidenced by a rise in the incidence in all tumour stages.[3,5] This concept is further validated by the observation of a stable rate of cases of renal cancer diagnosed only at autopsy over the last 40 years despite the more frequent use of imaging modalities.[6] The reasons for this rise in incidence are unknown.

From 1969 to 2002 the mortality rate for renal tumours in the United States and Canada increased.[2,3] Most of this rising trend in mortality seems to be accounted for by a significant increase until 1992. Mortality data over the last decade published in the USA (1992-2002) shows that mortality rates may have stabilized and in fact it appears that they are starting to decline.[1,2,7] This could be explained by the positive impact of the management of metastatic RCC or a true impact on mortality due to early detection and treatment. To date, there is no evidence to support one hypothesis over the other.

FIGURE 1. AGE-STANDARDIZED INCIDENCE RATES FOR SELECTED* CANCERS, MALES, CANADA, 1983–2012[1]



**Cancers (both sexes combined) with a statistically significant change in incidence rate of at least 2% per year*

2.2. Histology of Renal Masses

Most solid renal masses are RCCs. The histologic subtypes of RCC are classified according to the Heidelberg classification, the UICC and AJCC guidelines.[8,9] They include clear cell, papillary, chromophobe, collecting duct and unclassified RCC subtypes. Several studies have demonstrated that chromophobe and papillary carcinomas have a better prognosis than clear cell type carcinomas. In addition, patients who present with papillary and chromophobe RCCs tend to have tumours of lower stage compared with patients who have clear cell RCC.[10] To date, the proportion of different subtypes of RCC have not changed supporting the hypothesis that that risk factors have not changed significantly over time and observed changes are real.

Historical series reported that approximately 80-90% of solid renal masses were RCCs.[11] Several more recent reports have shown that for small renal masses the incidence of benign histology after partial or radical nephrectomy is as high as 48%.[10] With such high rates of benign disease, there is clearly a need for more pre-operative indicators of pathology. It has also been demonstrated that elderly patients with small renal masses are up to 3.5 times more likely to have benign lesions than RCC.[12]

2.3. Current Treatment of Renal Masses

Once a renal mass is identified on imaging studies, a treatment decision is made. The only established curative treatment for RCC remains surgery with radical or partial nephrectomy. For the past 50 years, the accepted standard treatment for RCC has been radical nephrectomy. More recently, excellent cancer control has been obtained with partial nephrectomy which has become the standard treatment for most small renal masses. Among those who undergo partial nephrectomy, excellent five and ten year survival rates are accompanied by a low risk of local (3.2%) and distant (5.8%) recurrence with preservation of renal function in 98% of cases.[13]

The rates of partial and radical nephrectomy are rising. Although morbidity from nephrectomy and partial nephrectomy has decreased, it is still significant and reported to be between 11% and 40%.[13,14] In particular, older patients, with

significant comorbid disease are at increased risk of perioperative mortality and morbidity.[15] Despite this rising trend in the use of partial and radical nephrectomies, the mortality rate from RCC has not decreased. In fact, the mortality rate had been rising steadily until recently.[1] Possible explanations for the increase in detection with minimal change in mortality include: 1) a lead time bias due to earlier detection, resulting in increased incidence but not changing the lifetime mortality risk from RCC; 2) length time bias resulting in the diagnosis of indolent tumours, without altering the mortality rate from life threatening tumours; or 3) mortality improvement that has yet to be seen for these small tumours with a long natural history and a wider window of curability. If this is the case, a change in mortality may only be expected 10 to 15 years following the era of increased detection, and we are only starting to see this now.[1]

With the introduction of Active Surveillance, ablative treatments (Radiofrequency Ablation and Cryotherapy) and laparoscopic and open partial nephrectomy for the management of small renal masses, the pretreatment histologic diagnosis of these masses has become very relevant. The pretreatment ability to classify tumours into benign and malignant, would aid to decrease the occurrence of unnecessary complications in patients with benign masses treated surgically and would increase the confidence in the patient and physician to decide on an expectant approach.

Different approaches have been evaluated in order to try to decrease the rate of surgery for benign disease.

2.4. Role of mass biopsy

The traditional indications for biopsy of small renal masses are limited to cases with atypical images, a suspected secondary tumour or to establish the primary tumour pathology in the presence of metastases. The technique, risks and results of needle biopsy have improved significantly as a result of better image guidance and one-handed automated needles with echogenic and depth markings.[16] Masses <1cm can be biopsied, although biopsies of masses >1cm and particularly >2 cm are more likely to provide useful information. A high degree of accuracy can be achieved with respect to tissue sampling interpretation. This has been reported as high as 100% sensitivity and specificity and 90% accuracy.[16] Volpe et al., demonstrated in a cohort of 100 patients that 84 (84%) biopsies were diagnostic for a malignant (66) or a benign (18) tumour.[17] Larger tumour size and a solid pattern were significant predictors of a diagnostic result. Histological subtyping and grading were possible on core biopsies in 93% and 68% of renal cell carcinomas, respectively. In that cohort, 20 patients underwent surgery after a diagnostic biopsy demonstrating a histological concordance of biopsies and surgical specimens was 100%.

Needle biopsy of a renal mass presumed to be RCC appears to be safe. A

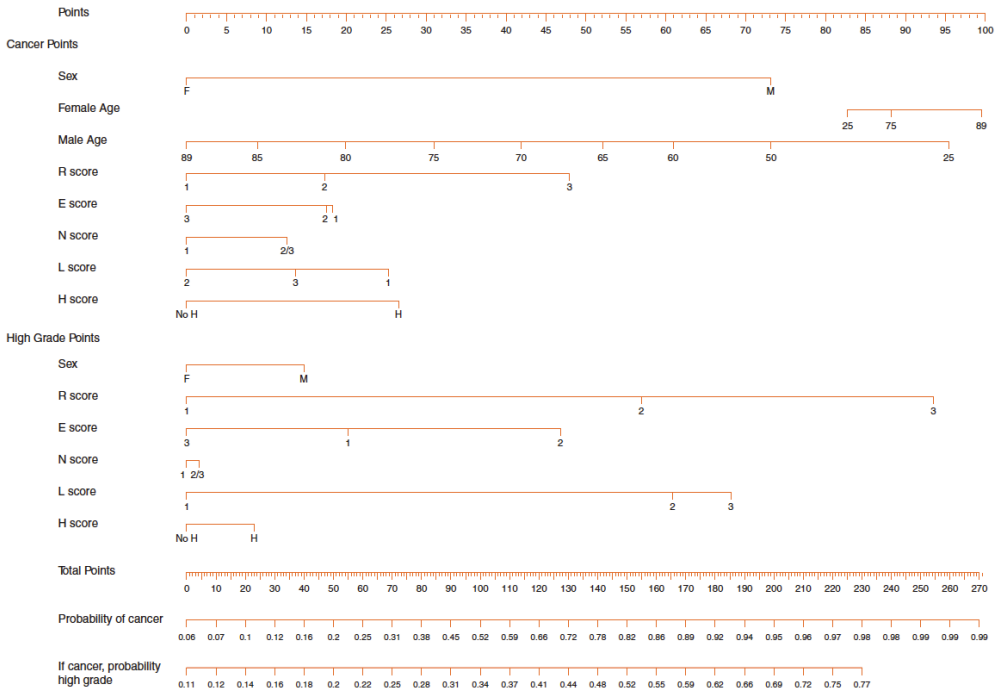
systematic review of the literature published on renal biopsy showed that in contemporary series, minor complications are rare (<5%), and catastrophic complications and mortality (no reported cases) are exceedingly rare.[18] While some degree of bleeding was evident on 85% to 91% of computed tomography (CT) images obtained routinely following renal mass biopsy, renal hemorrhage necessitating hospital admission or blood transfusion occurred in only 1-2% of cases and kidney loss was extremely rare. Other complications such as clinically significant pneumothorax were also found to be rare (<1%). Tumour seeding, a feared complication of renal mass biopsies, has a reported incidence of <0.01% and since 1994, no cases have been reported.

In spite of improved technique resulting in higher tissue yield and diagnostic accuracy as well as a decreased complication rate, the technical failure and indeterminate or inaccurate pathological diagnosis is still significant.[18] In addition, further studies that address the impact of tumour heterogeneity and sampling error for contemporary percutaneous renal biopsies are lacking. Furthermore, the amount of diagnostic and prognostic information obtained from histopathological analysis continues to be limited. Based on these facts, currently relatively young and healthy patients who are unwilling to accept the uncertainty associated with renal biopsies are still best treated with surgical excision. Therefore, a more rational approach to biopsying none or all renal masses, would be to biopsy all those masses with a higher risk of harbouring benign disease.

2.5. Tumour Morphology and Anatomy

Tumour morphology and anatomy also provide prognostic information regarding clinical behavior. It has been found that small, well marginated, and homogeneous tumours have lower nuclear grade, while higher grade lesions demonstrate irregular margins and greater inhomogeneity.[19] In a retrospective analysis of 193 patients, Zhang et al. found that clear cell RCC commonly (88–79%) and predictably (OR 22–54 in comparison with papillary/chromophobe tumours) presented with a mixed enhancement pattern of both hypervascular soft-tissue components and low-attenuation areas that corresponded to necrotic or cystic changes.[20] Kutikov et al. reported on the use of the R.E.N.A.L nephrometry scoring system to evaluate whether radiographic features correlated with histology and high-grade disease (Figure 2.).[21] The R.E.N.A.L. score was initially developed to standardize radiographic tumour reporting for assessment of perioperative complications and partial nephrectomy utilization rates. These authors found that nephrometry score correlated with both histology ($P<0.0001$) and grade ($P<0.0001$). They created a nomogram incorporating age, sex, and R.E.N.A.L. score with an area under the curve of 0.76 for histology and 0.73 for grade. Of importance, the R.E.N.A.L. component of the score only marginally added to the value of the nomogram, with size being the dominant component. Unfortunately these and other nomograms are rarely employed in clinical practice due to their complexity.

FIGURE 2. NOMOGRAM EVALUATING RISKS OF AN ENHANCING RENAL MASS BEING MALIGNANT AND HIGH GRADE.[21]



Total point values are independently calculated for the cancer and the high-grade models and then applied to the corresponding probability scale at the bottom of the figure

2.6. Prognostic Factors

2.6.1. Tumour Size

Renal masses can now be more accurately characterized by imaging using CT and magnetic resonance imaging (MRI). However specificity decreases to 80% for tumours ≤ 3.5 cm. At present, the only non-invasive and readily available preoperative prognostic factor is tumour size. As early as 1938, the association between tumour size and metastatic potential was recognized.[22] Many authors have demonstrated that a 4cm cut-off point accurately

differentiates prognosis and that patients who are treated for tumours this size or smaller have a local recurrence-free survival of 98.9%, distant recurrence-free survival of 97.8% and cancer-specific survival of 98.5%.[10,13,14,23,24] These figures are significantly better than those of larger tumours, with recurrence-free survival in the range of 40 to 90%.[10,23]

Studies evaluating the accuracy and reproducibility of CT image measurements of small renal masses have been performed and inter and intra-observer variability has been evaluated.[25] An excellent overall reliability coefficient of 97% in the X and Y axes mass diameters, and 96% in the Z axis mass diameter has been reported. This suggests that the measurement of renal tumour size on imaging is both accurate and reproducible.

Based on local observations, the location of renal masses appears to have a role in the distinction between benign and malignant masses. We have reported that central masses are five times more likely to be malignant when compared with peripheral ones.[26] Three additional groups have reported on the implications of tumour location in relation with pathological features. Frank and coauthors reported their series of outcomes in patients who underwent laparoscopic partial nephrectomies in 2006.[27] They included 154 patients with central tumours and 209 patients with peripheral masses. The average tumour size as measured by CT was 3.0 cm (range 1.0–7.0) and 2.4 cm (range 0.7–10.0) for central and peripheral masses respectively ($p<0.001$) In their

secondary analysis, they identified RCC in 80.1% and 65.5% of the patients with central and peripheral tumours respectively (p=0.002).

Venkatesh and colleagues reported in 2006 their outcomes of laparoscopic partial nephrectomies in 123 patients.[28] The average tumour size was 2.6 cm for all masses. In their series, only 55% of the peripheral renal masses had malignant histopathologic features, compared with 86%, 75% and 85.7% of the mesophytic, endophytic and hilar masses, respectively ($p<0.05$). They also found that the proportion of tumours having worse Fuhrman nuclear features (3 and 4) was only 3.7% in the exophytic masses, compared with 13.9% in the mesophytic and 25% in the endophytic lesions ($p<0.05$). Fuhrman grade is a scale from 1 to 4 that describes the grade of nuclear differentiation. Higher Fuhrman grades are associated with worse prognosis.

Nadu and coauthors also reported on the outcomes of laparoscopic partial nephrectomy performed in 212 patients.[29] In their series, 53 and 159 patients had central and peripheral masses, respectively. The reported proportion of malignant disease was 94% in the central masses compared with 82% of the peripheral ones. This difference was not found to be statistically significant.

2.6.2. Histology

Histological prognostic factors have been described after surgical removal of tumours. As discussed above, papillary and chromophobe tumours have more

benign behaviour than the other types of RCC. Presence of tumour necrosis has been associated with death from RCC in patients with clear cell and chromophobe subtypes but not papillary tumours.[30,31] Sarcomatoid differentiation has been proposed as a prognostic factor, but this finding is typically rare and not all studies have identified it as an independent prognostic feature.[31]

Nuclear grading has been associated with prognosis. The most commonly used grading system was developed by Fuhrman.[32] Current data supports the utility of nuclear grading in particular for papillary and clear cell RCC. Its applicability to other less frequent histological subtypes is not conclusive.

In summary, the incidence of small, incidentally detected renal masses presumed to be renal cell carcinoma is rising. The only treatment for RCC associated with decreased mortality is surgical removal with partial or radical nephrectomy. These procedures have a significant complication rate and affect the renal function. Traditionally, newly diagnosed renal masses are surgically removed based on image findings only and are rarely biopsied. Upon removal of these renal masses, up to 48% of them have proven to be benign, thereby exposing these patients to unnecessary risks. Currently, there are few pre-operative predictors of benign or malignant disease for these radiologically detected renal masses and these predictive tools have either low predictive ability or are

cumbersome to use or both. A classification tree method is proposed herein in order to predict benign vs. malignant disease.

2.7. Statistical Background

Prediction of histology based on patient and tumour characteristics has traditionally been based solely on tumour size. More recently, it has become clear that this prediction with more variables is a new and complex field. In the particular case of renal masses, relationships between variables are largely unknown and may be strongly nonlinear and involve high-order interactions. The commonly used exploratory and statistical modeling techniques often fail to find meaningful predictive patterns from such data. Classification and regression trees are statistical techniques ideally suited for both exploring and modeling this type of data.[33] As a non-parametric statistical method, classification trees are conventionally constructed for classification purposes and risk factor analyses.[34]

Trees explain variation of a single response variable based on one or more explanatory variables.[35] The response variable can be either categorical (classification trees) or numeric (regression trees), and the explanatory variables can be categorical and or numeric.

For classification purposes in medical literature, most reported studies show that there is really very little difference between the performance of logistic model methods and trees.[36-38] The rules and subgroups that are derived from trees are

promoted as being very easily described to clinicians since they logically demarcate clusters of symptoms, signs and other features.

The tree is constructed by continuously splitting the data, defined by a simple rule based on a single explanatory variable. At each split, the 'parent node' data is partitioned into two mutually exclusive and disjoint subsets ('children nodes'), each of which is as homogeneous as possible. The splitting procedure is then applied to each new group separately. The main objective is to partition the response into homogeneous groups, but also to keep the tree reasonably small. Splitting is continued until an overlarge tree is grown. The splitting procedure stops when all the records belong to the same class of response variable or all the records have identical attribute values (explanatory variables).

The way that explanatory variables are used to form splits depends on their type.[34] For a categorical explanatory variable with two levels, only one split is possible, with each level defining a group. For categorical variables with 3 or more levels, any combinations of levels can be used to form a split. For numeric explanatory variables, a split is defined by values less than, and greater than, a particular selected value. From all possible splits of all explanatory variables, the one that maximizes the homogeneity of the two resulting groups is selected.

During tree development, each subgroup (node) is typically characterized by the assignment of the majority class of the response variable (benign vs. malignant), group size (total number of benign and malignant lesions in that particular group),

and the values of the explanatory variables that define it (a cutoff point for continuous variables and the specific characteristic for categorical variables).

Trees are then represented graphically, with the root or parent node, which represents the original undivided data set, at the top, and the branches (internal nodes) and leaves (each leaf represents one of the final groups) beneath (Figure 5.).[34,35] Additional information can be displayed on the tree, e.g., summary statistics of nodes.

To determine how well a characteristic (explanatory variable) performs, the degree of impurity of the parent node (before splitting) and the degree of impurity of the children nodes (after splitting) need to be compared. The larger their difference, the better the predictive ability. The gain is a criterion that can be used to determine the goodness of a particular split. This measurement of impurity takes the value zero for completely homogeneous nodes (null tree), and increases as homogeneity decreases. Thus maximizing the homogeneity of the groups is equivalent to minimizing their impurity. Homogeneity can be defined in many ways, with the choice depending on the type of response variable. There are five measures of impurity (splitting criteria).[33] In this particular analysis, the Gini index is employed.[39] This index measures the heterogeneity in the sample. The formula for the Gini index is

$$i(p) = \sum_{i \neq j} p_i p_j = 1 - \sum_j p_j^2$$

Once the impurity has been measured in a node, the node splits and the drop in

impurity achieved by the split of the parent node into children nodes is a measure of the goodness of split. The formula for the goodness of split criterion is

$$\Delta i\,(s,N) = i(N) - p(N_L)\,i(N_L) - p(N_R)\,i(N_R)$$

where $i(N)$ is the value of the parent impurity, $p(N_R)$ is the probability of a case falling in the right daughter node and $p(N_L)$ is the probability of a case falling in the left daughter node.

This measure in drop of impurity is described as 'improve' in the rpart software package. Improve is a numerical value assigned to each explanatory variable. The variable with the highest improve is then chosen as the splitting criterion.

Once the tree has been developed, an evaluation of the performance of the classification model is based on the counts of outcomes correctly and incorrectly predicted by the model.[34,35] These counts are tabulated in a table known as a confusion matrix (Table 5.). Although a confusion matrix provides the information needed to determine how well a classification model performs, summarizing this information with a single number makes it more convenient to compare the performance of different models. This can be done using accuracy (number of correct predictions / total number of predictions) or can also be expressed in terms of its error rate (number of wrong predictions / total number of predictions). After the evaluation of the performance of the tree, an internal validation is performed. There are several internal validation methods.[40] In the Hold out method, the analyst assigns a proportion of the data (typically 50/50 or 70/30) for training (create the tree) and another for testing (internal validation). [34,35]The

main drawback with this method is that it does not use all the data for the creation of the tree. This is a particularly limiting step for small data sets or data sets with low frequency outcomes. The Random subsampling method is essentially when the Hold out method is repeated several times in order to try and improve the estimation of the performance.

The Cross-Validation method creates the tree in a large proportion of the population and leaves out a small proportion for testing.[34,35] This is performed systematically until all data has been used for training and testing the model. One version of the method carries out a 10-fold cross-validation where the data is divided into 10 subsets of equal size (at random) and then the tree is grown leaving out one of the subsets and the performance assessed on the subset left out from growing the tree. This is done for each of the 10 sets. The average performance is then assessed. The advantage of this method is that it uses the entire data set to train and test the model.

Most classification prediction models originating from complex data sets result in an overgrown tree that is difficult to use clinically and can suffer from data fragmentation. As recursive partitioning is a top-down approach, the number of records in each subgroup become smaller as we navigate down the tree. At the terminal nodes, the number of records may be too small to make a statistically robust prediction. For these reasons, most trees require subsequent pruning.

Pruning of the classification tree can be performed a priori by applying a set of early stopping rules to limit tree growth at inception: 1) stopping rule that dictates that a split must exceed a preset percentage of root node impurity; 2) ruling that splitting stops if a node contains less than 5 observations; or 3) stopping splitting if a node is pure or all observations are identical.[34,35]

Error-complexity pruning of the classification tree can also be performed after the tree is formed by trimming the fully grown tree in a bottom-up fashion to generate a sequence of pruned subtrees.[33] Cross-validation is used to estimate the overall impurity of each of the pruned subtrees and the tree with the lowest estimated overall impurity is selected.

Pruning of the tree can also be performed using a clinically based decision that allows the clinician to develop a tree that 'makes sense' for a particular use. If this approach is utilized, the performance of the pruned tree is measured using the confusion matrix and the overall accuracy is compared with that of the full tree.

Post-pruning tends to give better results than prepruning because it makes pruning decisions based on a fully-grown tree, unlike prepruning, which can suffer from premature termination of the tree-growing process.

The advantages of classification trees are that: 1) they can cope with any data structure or type; 2) they have a simple form such as a flow chart with great appeal for clinical use; 3) they use conditional information effectively; 4) they are invariant under transformations of the variables; 5) they are robust with respect to outliers; and 6) they give an estimate of the misclassification rate. The main disadvantages of classification trees are that: 1) they do not use combinations of variables as predictors; 2) trees can be deceptive – if a variable is not included it might appear to be “masked” by another variable; and 3) tree structures may be unstable – a change in the sample may give different trees, although this is also true for other modelling techniques.

2.8. Summary

The incidence of small incidentally detected renal masses is growing mostly at the expense of presentation in the elderly years when comorbidities are highly prevalent. The traditional approach for the treatment of renal masses is with surgical excision with partial or radical nephrectomy. These procedures carry a significant risk or complications. This is of particular importance, as up to 40% of these masses can be benign.

One approach to decreasing this high risk of benign disease in histological evaluation is to biopsy all renal masses. It has been demonstrated that with improved biopsy techniques, complication rates can be reduced, but the

information obtained by biopsies is still limited. Another approach is to develop predictive tools using patient and tumour characteristics. To date, only one group has reported on the prediction ability of their nomogram. This nomogram is complex and because of this, it is rarely used in clinical practice.

A classification tree was developed in order to aid in the pre-operative assessment of which patients have an increased risk of harbouring a benign renal mass and thereby decreasing the chances of offering an unnecessary operation to those patients who do not require it.

Chapter 3: Methods

3.1 Study Population / Sample

3.1.1. Inclusion Criteria:

All patients > 18 years of age who underwent a partial or radical nephrectomy for a renal mass < 5 cm in largest diameter presumed to be a renal cell carcinoma performed by Dr. Ricardo A. Rendon (RAR) and Dr. Joseph G. Lawen (JGL) at the Queen Elizabeth II Health Sciences Centre, Halifax, Nova Scotia, Canada between July 01, 2001 and June 30, 2010 were identified. These two surgeons were chosen as they performed the largest volume of these surgical procedures. The specified time period was selected in order to include contemporary cases with available imaging studies.

Dr. Lawen's cases were identified via billing codes for partial and radical nephrectomy. Dr. Rendon's cases were identified through a prospectively maintained database. Dr. Rendon's billings were also obtained in order to corroborate that no cases were missing in the database. Once all patients who had a partial or radical nephrectomy performed by the two surgeons during the specified period of time were identified, a review of all pathology reports was performed. All patients with masses that measured up to six centimeters

postoperatively were recorded. All preoperative abdominal imaging studies (CT, MRI, and ultrasound) were reviewed and patients with renal masses with a maximum one-dimensional diameter of less than five centimeters were ultimately included in the study. Only patients who had available information on all outcomes and potential predictor variables were included in the study.

Approval was obtained from the Capital District Health Authority Research Ethics Board.

3.1.2. Exclusion Criteria:

Patients undergoing a surgical intervention for a previously known benign tumour (large or symptomatic angiomyolipomas)

3.2. Data Collection

The following measures were abstracted from the prospectively maintained database (RAR's patients) and from a retrospective chart review (RAR's and JGL's patients) for all cases meeting inclusion and exclusion criteria using the data collection tool (Appendix 1).

3.3. Outcome of Interest:

Benign or malignant disease

Post-operative histological diagnosis was utilized to determine whether the excised mass was malignant or benign. Histologic subtypes for both benign and malignant lesions were collected.

3.4. Potential prognostic factors:

3.4.1. Tumour size:

Tumour size was obtained from the last CT or MRI performed pre-operatively. The maximum diameter in centimeters of each mass was recorded in all patients. When possible, two or three diameters in different dimensions were measured and recorded. All digital images were blindly (reviewer unaware of outcome of interest) and separately reviewed by Ross Mason (medical student) and RAR. Conflicts were resolved by re-review of images by both reviewers.

Tumour volume was calculated in one of three ways depending on the number of available dimensions: 1) for three dimensions, the formula for ellipsoid volume was employed (0.5326xyz), 2) for two different dimensions, the formula 0.5326xy(x+y/2) was used, and 3) for one dimension, the formula for volume of a sphere was employed ((4/3)(3.14)(x/2)^3) which is equivalent to 0.5326x^3.

3.4.2. Tumour location:

Central-Hilar-Peripheral

3.4.2.1. Central - A central mass was defined as one which extends into the kidney in direct contact with or invading into the collecting system and/or renal sinus (Figure 3.)

3.4.2.2. Hilar - A hilar mass was defined as one which is directly against or invading the main renal vessels (Figure 4.)

3.4.2.3. Peripheral - All other masses were defined as peripheral (Figure 5.)

FIGURE 3. CENTRAL RENAL MASS

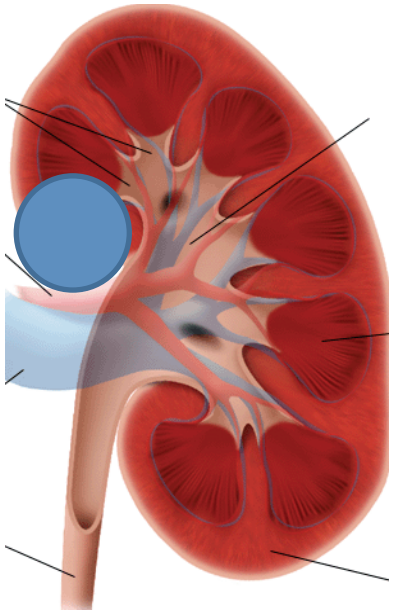


FIGURE 4. HILAR RENAL MASS

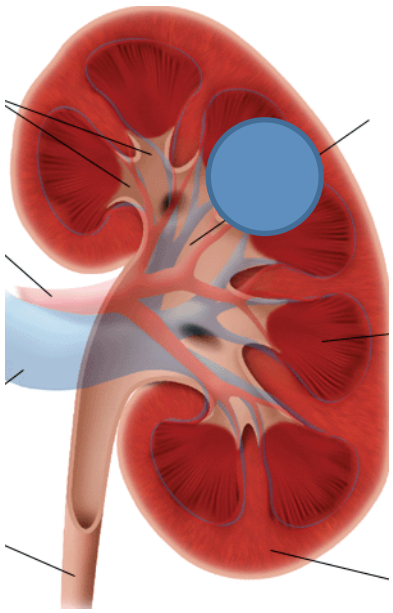
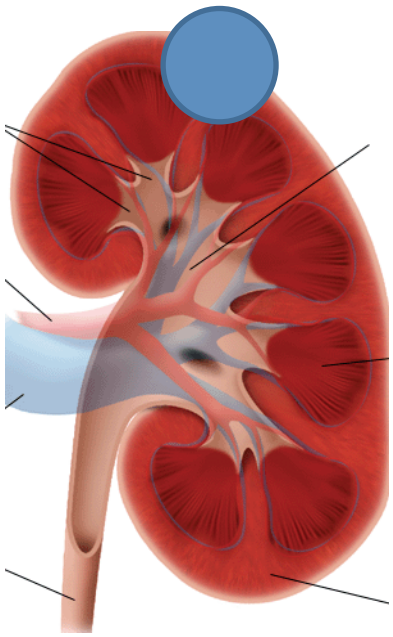


FIGURE 5. PERIPHERAL RENAL MASS



3.4.3. Degree of exophytic component:

After review of the images, all masses were assigned a percentage of exophytic component. This was defined as the amount of growth of the mass extruding beyond the contour of the kidney (1-100%).

3.4.4. Age in years

3.4.5. Sex

3.4.6. Symptoms at presentation:

Incidental - No symptoms at diagnosis

Hematuria – Microscopic or gross hematuria

Pain – Pain present at diagnosis and perceived by treating physician to be caused by the renal mass

As most cases in the current study presented with incidental tumours, it was decided to use a binary classification. Patients who presented with hematuria and pain were grouped into presentation with symptoms and all others were classified as incidental.

3.5. Statistical Analysis

3.5.1. Descriptive Statistics

Data preparation was performed in Excel (Microsoft Excel for Mac Version 14.2.3). Descriptive statistics were utilized to describe the study population. Means were employed to summarize continuous measures such as age,

tumour size and degree of exophytic component. Proportions were used to summarize categorical variables including final histology, symptoms at presentation, tumour location and Fuhrman grade.

3.5.2. Analytic Statistics

The main research question was addressed using Classification Tree methodology. Using the rpart package in the R language for statistical computing, a classification tree model was built using malignancy as the outcome variable and potential predictor variables as the covariates.[41] The covariates used were: age, symptoms at presentation, sex, degree of endophytic component, tumour volume, and central vs. peripheral location. The tree model was then validated using the 10-fold cross-validation routine in rpart, and all the error rate estimates were calculated for the 10 samples and then the mean estimate of the error rate was calculated.

The tree was pruned in order to develop the most parsimonious tree that made clinical sense (simpler) without a significant change in the predictive ability (overall accuracy). A confusion matrix was generated in order to obtain overall accuracy, sensitivity, specificity, positive and negative predictive values for the unpruned and pruned trees using malignancy as the outcome of interest. For the purposes of showing the results in a traditional framework a dummy variable with 5 levels was created. These levels correspond to the 5 terminal

nodes of the final pruned tree. The variable was then used as a predictor in a logistic regression model, the outcome being tumour histology. Odds ratios with 95% confidence intervals were calculated using as the reference the terminal node in the tree that corresponded to: Tumour volume > 5.67 cm^3, Endophytic component > 35% and Incidental diagnosis.

Analysis sequence, R code and SAS code are displayed in Appendices 2, 3 and 4.

Chapter 4: Results

In total 395 patients who fulfilled all the inclusion and exclusion criteria were identified.

4.1. Descriptive Results

Table 1 depicts patient characteristics. The median age was 61 years (range 24-90). Fifty five percent of the patients were male and 81% of the masses were detected incidentally, while 19% presented with symptoms such as hematuria or pain.

Table 1. Patient Characteristics

Median Age in years (range)	61 (24-90)
Sex	
Male	217 (55%)
Female	178 (45%)
Presentation	
Incidental	320 (81%)
Symptoms	75 (19%)

Table 2 depicts preoperative tumour characteristics. The median diameter of dimension X in CT or MRI was 3.1 cm (range 1-4.9), the median diameter of dimension Y was 2.8 cm (range 1.2-4.7), and the median diameter of dimension Z was 2.6 cm (range 1.1-4.7). For dimensions Y and Z there were 265 and 316

measurements missing respectively. The median tumour volume was 14.38 cm^3 (range 0.53-62.66). Two hundred and fifty two tumours (63.8%) had a central location and 143 (36.2%) were peripherally located.

Table 2. Preoperative Tumour Characteristics

Median Diameter in cm (range)		**Missing Values**
Dimension X	3.1 (1-4.9)	
Dimension Y	2.8 (1.2-4.7)	265
Dimension Z	2.6 (1.1-4.7)	316
MedianTumour Volume in cm^3 (range)	14.38 (0.53-62.66)	
Tumour Location		
Central	252 (63.8%)	
Peripheral	143 (36.2%)	

cm = centimeters
cm^3 = cubic centimeters

Table 3 displays the characteristics of the surgical intervention. Most (76.5%) had a laparoscopic procedure while only 23.5% had an open surgical approach. There were 223 patients (56.5%) who had a partial nephrectomy and 172 (43.5%) who had a radical nephrectomy.

Table 3. Treatment Characteristics

Surgical Procedure	N (%)
Laparoscopic	302 (76.5%)
Open	92 (23.5%)
Partial Nephrectomy	223 (56.5%)
Radical Nephrectomy	172 (43.5%)

N = number of observations

The median tumour diameter for dimension X as measured in the surgical specimen (pathology) was 3 cm (range 0.4-6.5), 3.5 cm (range 0.4-5.5) for dimension Y and 2.4 cm (range 0.4-5.5) (Table 4.). There were 2, 87 and 103 missing measurements for dimensions X, Y and Y respectively. Of the 395 patients with renal masses, there were 45 (10.4%) benign and 350 (89.6%) malignant lesions. The most common malignant subtype was Conventional Clear Cell Carcinoma in 264 (66.8%) masses. The predominant benign histology was Oncocytoma in 20 (5.1%) lesions. Amongst the benign tumours, the most common Fuhrman grade was 2 (51.5%), followed by Fuhrman 3 (24.8%), for a total of 81.4% of the masses with Fuhrman grades 2-3. Sixty one pathology reports did not describe the Fuhrman grade. The pathological stage of all 350 malignant lesions was available in 308 histological reports. Most (71.7%) patients were stage pT1a. Eighty-six (24.6%) patients were upstaged to pT1b and 14 (3.7%) to pT3a. Vascular and fat invasion was identified in 11 (2.8%) and 18 (4.6%) specimens respectively. The surgical margins were positive in 2.3% of the patients.

Table 4. Histopathological Findings

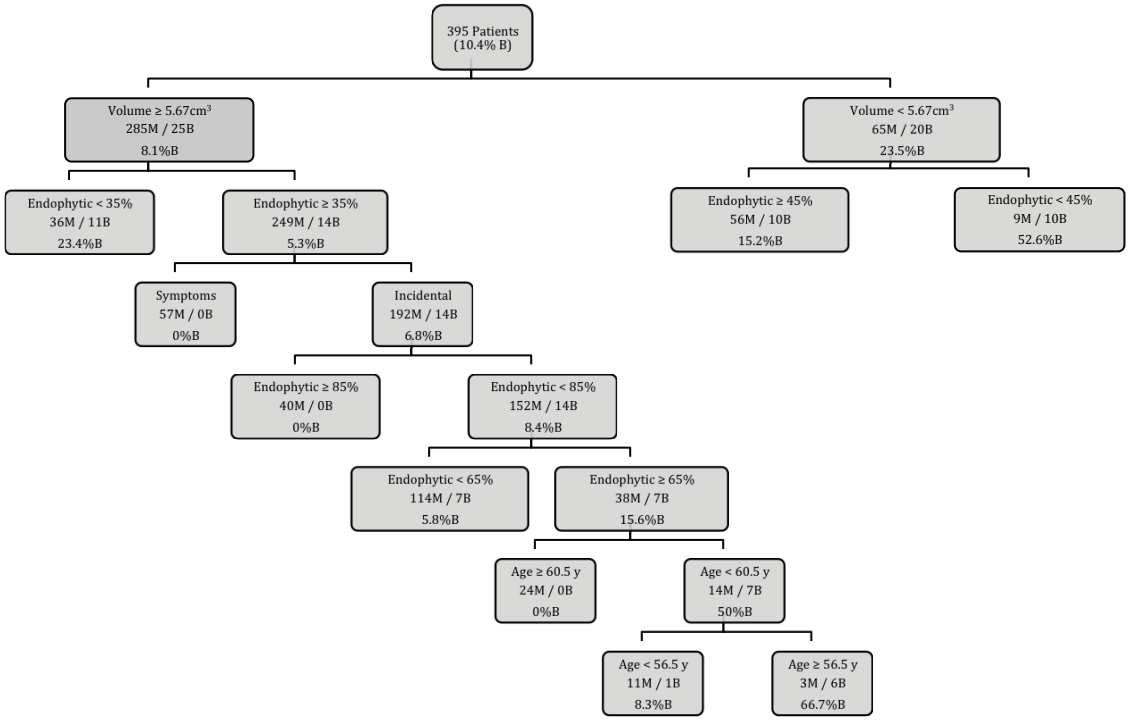
Tumour Characteristics		Missing Values
	Median cm (range)	
Diameter Dimension X	3 (0.4-6.5)	2
Diameter Dimension Y	3.5 (0.4-5.5)	87
Diameter Dimension Z	2.4 (0.4-5.5)	103
	N (%)	
Benign	45 (10.4%)	
Malignant	350 (89.6%)	
Histological Subtype		
Clear Cell	264 (66.8%)	
Papillary	68 (17.2%)	
Chromophobe	17 (4.3%)	
Granular Cell	1 (0.25%)	
Angiomyolipoma	11 (2.8%)	
Oncocytoma	20 (5.1%)	
Benign Cyst	6 (1.5%)	
Metanephric Adenoma	5 (1.3%)	
Leiomyoma	1 (0.25%)	
Cystic Nephroma	1 (0.25%)	
Spindle Cell	1 (0.25%)	
Vascular Invasion		
Yes	11 (2.8%)	
No	384 (97.2%)	
Fat Invasion		
Yes	18 (4.6%)	
No	377 (95.4%)	
Fuhrman Grade		61
1	39 (11.7%)	
1-2	7 (2.1%)	
2	172 (51.5%)	
2-3	17 (5.1%)	
3	83 (24.8%)	
3-4	1 (0.3%)	
4	15 (4.5%)	
Pathological Stage		42
1a	251 (71.7%)	
1b	86 (24.6%)	
3a	14 (3.7%)	
Margins		
Negative	386 (97.7%)	
Positive	9 (2.3%)	

cm = centimeters

4.2. Analytic Results

The graphic representation of the full classification tree can be seen in Figure 6; the results are depicted in a tabular format in Table 6. The model to predict benign vs. malignant histology from preoperative information was generated based on the following variables: age, symptoms at presentation, sex, degree of endophytic component, tumour volume, and central vs. peripheral location.

FIGURE 6. FULL CLASSIFICATION TREE



B= Benign
M = Malignant
y = years of age

The following variables, cutoff points and improvement measures were identified at the first node:

Volume	> 5.665266	improve = 3.1908580
Endophytic	< 35	improve = 3.0922840
Peripheral		improve = 3.0055210
Sex		improve = 1.2194190
Incidental		improve = 0.4135021

Where improve is the improvement in impurity given by a particular split.

From these results, it can be seen that even though the tree automatically splits using tumour volume, degree of endophytic component and peripheral vs. central location are almost as good predictors.

At the second node, the splitting happens based on < 35% degree of endophytic component:

Endophytic	< 35	improve = 2.6071720
Peripheral		improve = 1.4929760
Age	< 62.5	improve = 0.4622516
Sex		improve = 0.4439324
Volume	< 20.44918	improve = 0.4335753

Here it can be appreciated that there is a larger difference in the improvement measure between the first and second and third variables.

In the third node, the splitting occurs at < 45% degree of endophytic component:

Endophytic	< 45	improve = 4.14485400
Age	< 72.5	improve = 1.72376600
Volume	< 1.737608	improve = 0.86734220
Sex		improve = 0.78203370
Incidental		improve = 0.09490196

In the fourth node, the splitting occurs at symptoms at presentation:

Incidental		improve = 0.4124183
Endophytic	< 85	improve = 0.3673256
Peripheral		improve = 0.3430060
Volume	< 40.59451	improve = 0.2752715
Age	< 60.5	improve = 0.2689320

From this, it can be seen that the improvement measure for the first variable is lower than that observed in the previous nodes.

It can be appreciated that there is no further splitting on the right side of the tree after the third node and no additional splitting on the extreme left hand side of the left side of the tree.

The following are the variables, cutoff points and improvement measures that were observed for nodes 9, 19, 39, 79 (there were no additional splits at other nodes). These nodes correspond to split 5, 6, 7 and 8. It is important to mention

that the mathematical model creates a node for every single possible split, even though it is not statistically significant. For this reason, node 79 equates to split number 8. Moreover, nodes 40 to 78 were created but not displayed in the graph as they do not separate each subsample into subsequent different samples:

Node 9

Endophytic	< 85	improve = 0.45853320
Age	< 60.5	improve = 0.34951460
Peripheral		improve = 0.34919850
Volume	< 40.59451	improve = 0.32436010
Sex		improve = 0.01810746

Node 19

Endophytic	< 65	improve = 0.626249400
Age	< 62.5	improve = 0.487280100
Volume	< 40.59451	improve = 0.458698100
Peripheral		improve = 0.336415200
Sex		improve = 0.005497328

Node 39

Age	< 60.5	improve = 2.488889000
Volume	< 19.47239	improve = 1.043566000
Sex		improve = 0.012698410
Endophytic	< 75	improve = 0.006554019

Node 79

Age	< 56.5	improve = 3.500000
Volume	< 19.47239	improve = 1.121795

The calculated overall accuracy of the full tree is 89.6% with a 95% confidence interval of 86.2 to 92.5 (Table 5). The sensitivity is 96.6% and the specificity is 35.6%. The positive predictive value and the negative predictive value are 92.1% and 57.1% respectively. These performance measures were all calculated for malignancy as the main outcome.

Table 5. Classification Tree Performance

	Full Tree	Final Tree
Overall Accuracy (95% CI)	89.6% (86.2,92.5)	88.9% (85.3,91.8)
10-Fold Cross-validation Overall accuracy	85.4%	
Sensitivity*	96.6%	97.4%
Specificity*	35.6%	22.2%
Positive Predictive Value*	92.1%	90.7%
Negative Predictive Value*	57.1%	52.6%

CI = Confidence Interval
**Determined for Malignancy*

In order to internally validate the predictive tree, a 10-fold cross validation was performed. The 10 fold estimates of error rate were,

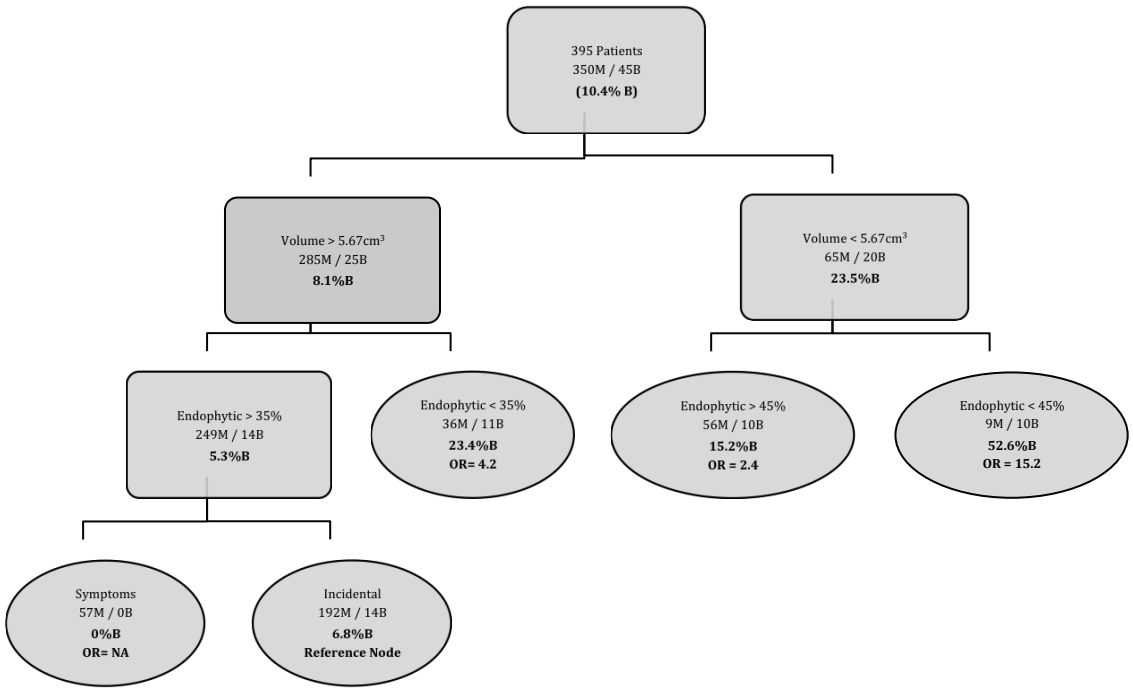
Fold	Actual Error Rate for Each Fold
1	0.23076923
2	0.17948718
3	0.17948718
4	0.12820513
5	0.20512821
6	0.12820513
7	0.12820513
8	0.02564103
9	0.10256410
10	0.15384615

The mean estimate error rate of the 10 folds was 0.1461538, which results in an overall predictive accuracy (accuracy = 1 – error) of 0.8538462 or 85.4%. As expected, the overall 10-fold cross-validated error rate, is slightly higher than that generated by re-substitution error rate but still at a clinically acceptable level.

This analysis resulted in a rather complex tree that is difficult to use in clinical practice. When we tried to perform error-complexity pruning, the low frequency of event (benign) resulted in the generation of either a very large, non-clinically useful tree or a null tree (parent node only) with only one terminal leaf. Based on this, we employed a clinically based pruning of the full tree. It can be seen that after node 3, all subsequent partitions originated from one of the two partitions of

node 2. This resulted in the pruned tree illustrated in Figure 7. Based on this final tree, 5 disjoint groups of patients with varied risks of benign disease can be depicted in a tabular form (Table 6.).

FIGURE 7. FINAL CLASSIFICATION TREE



B= Benign
M = Malignant
y = years of age
OR = Odds Ratio

Oval shape depicts terminal leafs of the predictive tool

Table 6. Classification Tree – Terminal Nodes (Final Pruned Tree)

Terminal Node	N	Benign N (%)	OR* (95% CI)
Vol > 5.67 cm^3, Endo > 35%, Incidental	206	14(6.8)	Reference
Vol > 5.67 cm^3, Endo > 35%, Symptoms	57	0(0)	NA
Vol > 5.67 cm^3, Endo < 35%	47	11(23.4)	4.19 (1.76,9.97)
Vol < 5.67 cm^3, Endo > 45%	66	10(15.2)	2.45 (1.03,5.81)
Vol < 5.67 cm^3, Endo < 45%	19	10(52.6)	15.24 (5.33,43.61)

N = Number of observations in the node
OR = Odds Ratio
CI = Confidence interval
Vol = Tumour volume at diagnosis
Endo = Degree of endophytic component
** Results from Logistic Regression Analysis*

The calculated overall accuracy of the final pruned tree is 88.9% with a 95% confidence interval of 85.3 to 91.8 (Table 5). The sensitivity is 97.4% and the specificity is 22.2%. The positive predictive value and the negative predictive value are 90.7% and 52.6% respectively. This overall accuracy of the pruned tree compares favourably with that calculated for the full tree. This demonstrates that by simplifying and creating a more clinically useful decision tree, there is no significant decrease in model performance.

Chapter Five: Discussion

Small renal masses are becoming a growing problem. The incidence is rising mostly at the expense of incidental findings in the elderly. Over the past decade we have learned that not all these small masses are in fact renal cell carcinoma and more importantly, even those that are RCCs may not need to be treated. In addition, the treatment of these masses is complex and often leads to significant morbidity as a result of immediate postoperative complications or due to loss of renal mass and subsequent renal insufficiency. Based on these facts, it is becoming increasingly clear that the traditional approach for the management of small renal tumours, with immediate surgical intervention, results in overtreatment and in many cases, unnecessary morbidity. In order to overcome this problem, two areas of knowledge need to be improved: 1) The nature (histology) of renal masses and 2) The natural history of untreated renal masses.

In this study the nature of renal masses was explored. The predictive ability of patient and tumour characteristics to differentiate between benign and malignant histology was explored. This study included 395 patients who underwent laparoscopic or open partial or radical nephrectomies for renal masses less than 5 cm in largest diameter. The age and size distribution represent that of a surgical cohort. The present series is populated by a larger number of endophytic and central renal masses when compared to other large series reported in the literature. This is the result of a referral bias of one of the surgeons (RAR) whose

practice is based on the management of more complex renal masses treated with partial nephrectomy for a large catchment area. Even though this may introduce a selection bias in this population, the cohort was partially balanced by inclusion of other cases more reflective of a general practice (JGL's patients).

The rationale for selection of masses smaller than 5 cm in largest diameter is the fact that larger masses start to engulf more of the kidney and therefore the proportion of peripheral masses decreases dramatically. One of the problems with all predictive tools is that for rare outcomes, larger samples are necessary. In the present study, close to 11% of the masses were benign. The inclusion of larger masses, which have a lower likelihood of harbouring benign disease, would have resulted in a lower proportion of benign disease. Even though these size inclusion criteria may appear restrictive, it in fact represents close to 70% of all renal masses diagnosed nowadays. Based on all these facts, we believe that this inclusion criteria represents well the population where the main management dilemma exists.

In the past, several authors, including our group, have reported on isolated patient and tumour characteristics as predictors of benign disease. Regarding patient characteristics, several authors have reported on the relationship between sex and histology. In 2007, Lane et al. reported their findings on preoperative patient characteristics as a predictor of benign versus malignant histology and found that younger women were more likely to harbour benign

disease (36%) than men (8%); on the other hand, older men had a higher likelihood of benign disease than older women.[42] Hsieh et al. demonstrated that the incidence of benign tumour was greater in females (p=0.014) and tumour size 2 cm or less (p=0.02), compared with males and tumour size more than 2 cm, respectively.[43]

Regarding tumour characteristics, in a previous study, our group demonstrated that peripheral masses are more likely to be benign than central ones.[26] In this study we identified that the proportion of benign disease by location was 5.9% and 19.5% for central and peripheral masses, respectively. The effect of location was found to have a significant prognostic value (p = 0.0273) with an adjusted odds ratio of 3.51 (95% CI = 1.38-19.62) for the odds of a benign diagnosis in peripheral compared to central tumours. Tumour size and patient sex were not significant predictors of benign pathology (p = 0.483 and 0.191, respectively). Malignant histology has also been found to be strongly correlated with tumour size. Frank et al. found that 46% of tumours less than 1.0cm in maximal diameter were benign and only 2.3% of these tumours were high grade.[44] Conversely, only 6.3% of tumours ≥ 7 cm were benign and 57.7% were high grade. Clear cell RCC, which has a higher malignant potential than papillary RCC, was also more common in larger tumours. Thompson and his group reported in a large surgical dataset, that the odds ratio for the association of malignancy with tumour size was 1.16 (95% CI 1.11–1.22, P<0.001), suggesting that each incremental centimeter in tumour size was associated with a 16% increase in the odds of

malignancy.[45]

Instead of utilizing single patient and tumour parameters, and using several patient and tumour characteristics, the current study demonstrated that benign vs. malignant can be predicted with an overall accuracy of the full tree of 89.6% with a 95% confidence interval of 86.2 to 92.5. This is the first time a classification tree has been used to determine preoperatively whether a renal mass is benign or malignant.

Using this tree for example, a patient with a renal mass with a volume of < 5.66 cm^3 that is < 45% endophytic has a 52.6% chance of having benign pathology. Conversely, a renal mass with a volume > 5.66 cm^3 that is > 35 % endophytic has only a 5.3% possibility of being benign. More over, a patient with similar characteristics but who presented with symptoms at diagnosis has a 0% chance of having a benign tumour. The risk difference of harbouring malignant disease between these three hypothetical patients is clearly different and this finding can be of help when deciding which patients should undergo further evaluation before exposing them to morbid surgical procedures. From the results of the logistic regression analysis modeling benign diagnosis, it can be seen that there is an increasing trend from left to right in the odds of having a renal mass that harbours benign disease. Of note, the odds ratio estimates have wide confidence intervals due to the low event frequencies (45/395). The implication of these

findings is that not surprisingly, the logistic regression analysis demonstrates a predictive tool with generalizability limitations.

Other groups have taken different approaches to determine the risk of benign vs. malignant disease using multivariable models. Kutikov et al. reported on the use the R.E.N.A.L nephrometry scoring system to evaluate whether radiographic features correlated with histology and high-grade disease.[21] The R.E.N.A.L. score was initially developed to standardize radiographic tumour reporting for assessment of perioperative complications and partial nephrectomy utilization rates. These authors found that nephrometry score correlated with both histology ($P<0.0001$) and grade ($P<0.0001$). They created a nomogram incorporating age, sex, and R.E.N.A.L. score with an area under the curve of 0.76 for histology and 0.73 for grade (Figure. 2). Of importance, the component of the score only marginally added to the value of the nomogram, with size being the dominant component. Specifically, the p values for age, sex, age-sex interaction, R, E, N, L and H scores were 0.053, 0.189, 0.203, <0.001, 0.041, 0.951, 0.002 and 0.583 (R = radius, E = exophytic/endophytic; N = nearness to collecting system, L = location , H = hilar). From this we can see that one of the major drawbacks of using predictive models in other situations (prediction of histology) rather than its original purpose (prediction of perioperative complications) is that the model uses variables that are not predictive, thereby reducing the overall predictive ability of the tool.

Even though the RENAL score model has been made available as a Web tool for point-of-service use (www.cancernomograms.com), unfortunately this nomogram and other traditional regression models are rarely employed in clinical practice due to their complexity and the fact that they are hard to memorize and time consuming (Figure 2.). Conversely, a huge appeal of classification and regression tree (CART) methodology is that it makes a very intuitive diagram to represent risk. Such a diagram is very visually appealing and not difficult to memorize for the clinician. Furthermore, these intuitive diagrams replicate the clinician's thought process when encountering a diagnostic dilemma and using decision rules (trees).

A downside of CART is its predictive accuracy. In most cases, CART oversimplifies the risk determination. In Figure 7, it can be seen that the highest risk patients' need is tumour volume > 5.66 cm^3 only and nothing else matters. As detailed in the results section, two other variables were closely competing with tumour volume as the number one splitting factor: Volume < 5.66 (improve = 3.1908580), Endophytic component < 35 (improve = 3.0922840) and Peripheral vs. central component (improve = 3.0055210). From this it could be hypothesized that the model may be unstable at this node and with an increase in sample size, small variations on predictive ability of any of these variables, the node could split on a different characteristic, thereby changing the classification tree.

Another downside of the CART approach is that it assumes that those patients in a particular group are homogeneous with respect to risk. Some specific comparisons of CART and neural networks with traditional regression equations, which would not make the homogeneous risk assumption, tend to favor regression as a more accurate approach.[46]

A final downside of this predictive tool, as it is for every other predictive modeling tool, is data fragmentation. As recursive partitioning is a top-down approach, the number of records in each subgroup become smaller as we navigate down the tree. At the terminal nodes, the number of records may be too small to make statistically robust predictions.

This classification tree demonstrated a sensitivity of 96.6%, a specificity of 22.2%, with positive and negative predictive values of 92.1% and 57.1% respectively. This high sensitivity (proportion of true malignant lesions that are correctly identified by the predictive model) is very important when trying to differentiate benign vs. malignant histology. In oncology in general, high sensitivity is desired in order to decrease the chance of missing malignant cases. On the other hand, the specificity (proportion of true benign lesions that are correctly identified by the predictive model) is relatively low (22.2%). For this particular model this low specificity is a drawback. One way to deal with this problem is to adjust the pre-modeling probability of a lesion being benign by assigning more weight to this outcome.[35]

Finally, another way to identify the histologic nature of renal masses is to perform a percutaneous biopsy in all patients at the time of diagnosis. As discussed earlier, the use of percutaneous biopsy in the evaluation of renal masses has been limited by concerns of false negative results and complications. Over the last decade, imaging and biopsy techniques have improved, thereby increasing the accuracy in detecting malignancy and decreasing the false negative rate in experienced centres. Additionally, with the evolution of biopsy techniques and utilization of a core needle biopsy through a coaxial cannula, no cases of tumour seeding were reported in the last decade. Experienced centres have also reported complication rates that do not exceed 10%, consisting mainly of hematomas treated conservatively by bed rest and transfusion in the majority of cases. Other complications such as arteriovenous fistula, pneumothorax, and bowel perforation are now rare. It is important to mention that the published literature on this topic comes from a handful of centres, and to date, there is no data available on similar procedures performed by centres with less expertise.

Even though percutaneous biopsy of renal masses can now be safely performed with high sampling accuracy in experienced centres, we continue to depend on limited histological information provided by these biopsies. It has been demonstrated that biopsies tend to be adequate for differentiation between benign and malignant disease but have difficulties differentiating chromophobe RCC (malignant) from oncocytoma (benign). Moreover, further studies that address the impact of tumour heterogeneity within the same mass and sampling error for contemporary percutaneous renal biopsies are lacking. Furthermore, the

amount of diagnostic and prognostic information obtained from histopathological analysis continues to be limited to histologic subtype, which alone has not shown to provide prognostic information in any of the studies evaluating active surveillance for the management of renal masses. In the future, other histologic parameters and tissue molecular markers that have not permeated the clinical scene, afforded by tissue biopsy, may help formulate therapeutic plans for patients.

Based on the aforementioned facts, it seems that a blanket approach to biopsying all renal masses is unjustified and a more risk-based approach is necessary.

Clinical Impact

This classification tree for the prediction of benign vs. malignant histology of renal masses at diagnosis based on patient and tumour characteristics has great overall accuracy and sensitivity but low specificity. This predictive tool is a major improvement over the current treatment algorithm, which is to surgically excise or biopsy all renal masses. This classification tree identifies masses that have a high risk (> 50%) of benign disease, for which a biopsy is recommended. Furthermore, it identifies masses that have a high risk (0 – 5%) of malignant disease requiring surgical removal in the young and healthy individual. Based on the low (10%) frequency of the main outcome (benign disease) and the low specificity of this predictive tool, there is a significant proportion of masses that

are classified as intermediate risk for harbouring benign disease. Currently there are no better predictive tools than the one discussed in this study. In spite of the deficiencies of this predictive tool, it represents a significant improvement in the management of small renal masses. Further refinements to this predictive tool are underway.

Chapter Six: Conclusion

6.1. Conclusion

Preoperative tumour histology (benign vs. malignant disease) can be predicted using a classification tree based on patient and tumour characteristics with an 89.6% overall accuracy. This predictive accuracy is higher than that shown in biopsy series.[47] In this era of incidentalomas, tumour size maintains its predictive ability to differentiate benign vs. malignant lesions. Based on these data, it is hypothesized that for smaller renal masses that are less endophytic, a pre-operative biopsy may be indicated due to the higher risk of benign pathology. Conversely, larger and more endophytic masses have a very high risk of harbouring malignant disease and those should be surgically excised without further investigations. We believe that classification trees like the one proposed herein are easier to use in the clinical setting when compared with logistic regression models as they mimic the clinician's thought process and therefore will be more utilized.

6.2. Strengths and Limitations

The main strength of this study is its novelty. Although one other group has utilized a multivariable approach to differentiate malignant from benign disease in

this context, there are no reports on prediction of benign vs. malignant histology using a classification tree method.[21] Moreover, the nomogram used by Uzzo's group was developed for another purpose, thereby exposing the model to lower predictive accuracy when used to predict benign or malignant histology.

Another strength lies in the fact that most of the series that identify prognostic factors report information on patients accrued before this era of incidental diagnosis at earlier stages. As described above, to date, the most important pre-operative prognostic factor has been tumour size. Currently, the vast majority of renal masses are identified when they are still small in size and therefore we hypothesize that tumour size might not carry the prognostic weight it once did.

All clinical, radiographical and histological data have been analyzed retrospectively. In order to minimize any potential biases, the reviewers of the imaging studies were blinded to the outcomes of interest. Additionally, all sequential patients who underwent surgery during the study period were included in the study. By following these two steps, it was intended to minimize any potential biases introduced by retrospective data collection.

In this study, the most important predictor of benign disease was tumour volume. Even though in clinical practice it is customary to use a single diameter to measure tumour size, tumour volume provides a better representation of the real tumour size as not all masses are perfect spheres. The dataset utilized for this

analysis had many missing values for dimensions Y and Z, 265 and 316 respectively. Even though tumour volume can be derived from a single diameter, it is hypothesized that the three diameters would have provided a more accurate estimate of volume.

From a statistical point of view, the robustness of a Classification Tree (chances of classification errors) lies in the sample size and the frequency of the outcome variable. This particular dataset has an overall 11% event rate and event rates within the different levels of the predictive tool, ranged from 0 to 52.6%. These low event rates, make statistical modeling computationally challenging which is true for all statistical modeling approaches.

Finally, the 10-Fold Cross-validation overall accuracy for the final (pruned tree) has not been calculated. The rpart package determines this for the trees that are pruned using the automated cost-complexity pruning approach. As described above, the chosen tree was clinically (manually) pruned. In order to perform the 10-Fold Cross-validation this would have to be performed manually by creating 10 different subsamples and then calculate the error estimate for all and obtain the means of all error rates. These steps will be performed in a future analysis.

6.3. Future Directions

6.3.1. External Validation

Ethics approval has been obtained at Princess Margaret Hospital (PMH), Toronto, Ontario and CDHA, Halifax, Nova Scotia to validate the current predictive tool using an external sample.

6.3.2. Integration of Predictive Tools

It has become evident that in medicine, with few exceptions, a single approach to solve a problem tends to fail. It is our intent to combine our Classification Tree with data from a cohort of 500 patients who underwent renal biopsies for small renal masses at PMH. It is our opinion that neither the current Classification Tree alone or renal biopsies for all patients perform as well separately as when they are combined. It is hypothesized that using the Classification Tree to direct biopsies for masses with higher risk of harbouring benign pathology and immediate surgery for masses with high risk of malignancy will provide a lower risk and a more comprehensive approach.

References

1. Canadian Cancer Statistics 2012. In. Toronto, ON: Canadian Cancer Society; 2012.

2. Cancer Facts and Figures. American Cancer Society, 2005. (Accessed at http://www.cancer.org/downloads/STT/CAFF2005f4PWSecured.pdf.)

3. Chow WH, Devesa SS, Warren JL, Fraumeni JF, Jr. Rising incidence of renal cell cancer in the United States. JAMA 1999;281:1628-31.

4. UICC IUAC. AJCC Cancer Staging. 7th ed: Wiley-Liss; 2010.

5. Hock LM, Lynch J, Balaji KC. Increasing incidence of all stages of kidney cancer in the last 2 decades in the United States: an analysis of surveillance, epidemiology and end results program data. J Urol 2002; 167:57-60.

6. Mindrup S, Pierre J, Dahmoush L, Konety B. The Prevalence of Renal Cell Carcinoma Diagnosed at Autopsy. BJU Int 2005;95:31-3.

7. Canadian Cancer Statistics 2009. 2009. (Accessed at http://www.cancer.ca.)

8. Storkel S, Eble JN, Adlakha K, et al. Classification of renal cell carcinoma: Workgroup No. 1. Union Internationale Contre le Cancer (UICC) and the American Joint Committee on Cancer (AJCC). Cancer 1997;80:987-9.

9. Kovacs G, Akhtar M, Beckwith BJ, et al. The Heidelberg classification of renal cell tumours. J Pathol 1997;183:131-3.

10. Frank I, Blute ML, Cheville JC, Lohse CM, Weaver AL, Zincke H. Solid renal tumors: an analysis of pathological features related to tumor size. J Urol 2003;170:2217-20.

11. Kutikov A, Fossett LK, Ramchandani P, et al. Incidence of benign pathologic findings at partial nephrectomy for solitary renal mass presumed to be renal cell carcinoma on preoperative imaging. Urology 2006;68:737-40.

12. Sanchez-Ortiz R, Luongo T, Tamboli P, Madsen L, Swanson D, Wood C. Increased Incidence of Benign Histology in Elderly Patients with Renal Masses. J Urol 2005;173:381.

13. Novick AC. Safety and efficacy of partial nephrectomy for all T1 tumors based on an international multicenter experience. Patard JJ, Shvarts O, Lam JS, Pantuck AJ, Kim HL, Ficarra V, Cindolo L, Han KR, De La Taille A, Tostain J, Artibani W, Abbou CC, Lobel B, Chopin DK, Figlin RA, Mulders PF, Belldegrun AS, Department of Urology, University of California Los Angeles, Los Angeles, CA. Urol Oncol 2005;23:71-2.

14. Lee CT, Katz J, Shi W, Thaler HT, Reuter VE, Russo P. Surgical Management of renal tumors 4 cm. or less in a contemporary cohort. J Urol 2000;163:730-6.

15. Kemeny MM, Busch-Devereaux E, Merriam LT, O'Hea BJ. Cancer surgery in the elderly. Hematol Oncol Clin North Am 2000;14:169-92.

16. Caoili EM, Bude RO, Higgins EJ, Hoff DL, Nghiem HV. Evaluation of sonographically guided percutaneous core biopsy of renal masses. AJR Am J Roentgenol 2002;179:373-8.

17. Volpe A, Mattar K, Finelli A, et al. Contemporary results of percutaneous biopsy of 100 small renal masses: a single center experience. J Urol 2008;180:2333-7.

18. Lane BR, Samplaski MK, Herts BR, Zhou M, Novick AC, Campbell SC. Renal mass biopsy--a renaissance? J Urol 2008;179:20-7.

19. Birnbaum BA, Bosniak MA, Krinsky GA, Cheng D, Waisman J, Ambrosino MM. Renal cell carcinoma: correlation of CT findings with nuclear morphologic grading in 100 tumors. Abdom Imaging 1994;19:262-6.

20. Zhang J, Lefkowitz RA, Ishill NM, et al. Solid renal cortical tumors: differentiation with CT. Radiology 2007;244:494-504.

21. Kutikov A, Smaldone MC, Egleston BL, et al. Anatomic features of enhancing renal masses predict malignant and high-grade pathology: a preoperative nomogram using the RENAL Nephrometry score. Eur Urol 2011;60:241-8.

22. Bell ET. A Classification of Renal Tumours with Observations on the frequency of the Various Types. J Urol 1938;39:238.

23. Frank I, Leibovich BC, Cheville JC, Lohse CM, Zincke H, Blute ML. Preoperative Prediction of Pathologic Features and Outcome in Patients with Clinically Confined Solid Renal Tumors. J Urol 2005;173:260.

24. Krejci KG, Frank I, Blute ML, Weaver AL, Zincke H. Grade- and size-specific outcomes for stage T1 renal cell carcinoma (RCC) after nephron sparing surgery (NSS). J Urol 2001;165:A 651.

25. Punnen S, Haider MA, Lockwood G, Moulding F, O'Malley ME, Jewett MA. Variability in size measurement of renal masses smaller than 4 cm on computerized tomography. J Urol 2006;176:2386-90; discussion 90.

26. Mason RJ, Abdolell M, Rendon RA. Tumour location as a predictor of benign disease in the management of renal masses. Canadian Urological Association journal = Journal de l'Association des urologues du Canada 2010;4:414-7.

27. Frank I, Colombo JR, Jr., Rubinstein M, Desai M, Kaouk J, Gill IS. Laparoscopic partial nephrectomy for centrally located renal tumors. J Urol 2006;175:849-52.

28. Venkatesh R, Weld K, Ames CD, et al. Laparoscopic partial nephrectomy for renal masses: effect of tumor location. Urology 2006;67:1169-74; discussion 74.

29. Nadu A, Kleinmann N, Laufer M, Dotan Z, Winkler H, Ramon J. Laparoscopic partial nephrectomy for central tumors: analysis of perioperative outcomes and complications. J Urol 2009;181:42-7; discussion 7.

30. Frank I, Blute ML, Cheville JC, Lohse CM, Weaver AL, Zincke H. An outcome prediction model for patients with clear cell renal cell carcinoma treated with radical nephrectomy based on tumor stage, size, grade and necrosis: the SSIGN score. J Urol 2002;168:2395-400.

31. Lohse CM, Cheville JC. A review of prognostic pathologic features and algorithms for patients treated surgically for renal cell carcinoma. Clin Lab Med 2005;25:433-64.

32. Fuhrman SA, Lasky LC, Limas C. Prognostic significance of morphologic parameters in renal cell carcinoma. Am J Surg Pathol 1982;6:655-63.

33. Breiman L, Friedman J, Olshen R, Stone C. Classification and Regression Trees. Belmont, CA: Wadsworth; 1984.

34. Zhang H, Singer BH. Recursive Partitioning and Applications. 2nd ed: Springer; 2010.

35. Berk RA. Classification and Regression Trees (CART). In: Berk RA, ed. Statistical Learning From a Regression Perspective: Springer Science and Busines Media; 2008:103-67.

36. McGuire V, Nelson LM, Koepsell TD, Checkoway H, Longstreth WT, Jr. Assessment of occupational exposures in community-based case-control studies. Annual Review of Public Health 1998;19:35-53.

37. Nelson LM, Bloch DA, Longstreth WT, Jr., Shi H. Recursive partitioning for the identification of disease risk subgroups: a case-control study of subarachnoid hemorrhage. Journal of Clinical Epidemiology 1998;51:199-209.

38. Selker HP, Griffith JL, Patil S, Long WJ, D'Agostino RB. A comparison of performance of mathematical predictive methods for medical diagnosis: identifying acute cardiac ischemia among emergency department patients. Journal of investigative medicine : the official publication of the American Federation for Clinical Research 1995;43:468-76.

39. Kakwani N. Income Inequality and Poverty: Methods of Estimation and Policy Analysis. Washington and Oxford: World Bank and Oxford University Press; 1980.

40. Clark LA, Pregibon D. Tree-based models. In: Chambers JM, Hastie TJ, eds. Statistical models in S. Grove, California, USA: Wadsworth and Brooks, Pacific; 1992:377-420.

41. Team RDC. R: a language and environment for statistical computing. R foundation for statistical computing. In. Vienna, Austria; 2008.

42. Lane BR, Babineau D, Kattan MW, et al. A preoperative prognostic nomogram for solid enhancing renal tumors 7 cm or less amenable to partial nephrectomy. J Urol 2007;178:429-34.

43. Hsieh PF, Chang CH, Huang CP, et al. The impact of gender and size on the pathology of small renal mass. The Kaohsiung journal of medical sciences 2012;28:369-72.

44. Frank I, Blute ML, Cheville JC, Lohse CM, Weaver AL, Zincke H. Solid renal tumors: an analysis of pathological features related to tumor size. J Urol 2003;170:2217-20.

45. Thompson RH, Kurta JM, Kaag M, et al. Tumor size is associated with malignant potential in renal cell carcinoma cases. J Urol 2009;181:2033-6.

46. Kattan MW. Comparison of Cox regression with other methods for determining prediction models and nomograms. J Urol 2003;170:S6-9; discussion S10.

47. Volpe A, Kachura JR, Geddie WR, et al. Technique, safety and accuracy of sampling of renal tumors by fine needle aspirationand core biopsy. J Urol 2007;178:379-86.

Appendix 1. Data Collection Tool

Case ID: _____

Demographics:

Age	Sex

Symptoms at presentation:

Incidental	Hematuria	Pain	Other

Diagnostic imaging information:

Tumour size (cm):

A/P	Transverse	Coronal	Volume

Tumour location: C H P

Endophytic: _____%

Surgical information:

Laparoscopic ___ Open ___

Partial Nephrectomy ___ Radical Nephrectomy ___

Histological information:

Tumour type:

Clear cell	Papillary	Chromophobe	Angiomyolipoma	Oncocytoma	Other

Fuhrman grade: 1 2 3 4

Pathological Stage ___

Vascular Invasion: Yes ___ No ___

Fat Invasion: **Yes ____** **No ____**

Margins: **Positive ____** **Negative ____**

Tumour diameter:

Dimension x: ____
Dimension y: ____
Dimension z: ____

Appendix 2. Analytical Sequence

STEP #	ANALYTICAL SEQUENCE
1	rpart, caret and tree packages are loaded in R
2	The dataset is duplicated and renamed
3	The descriptive summary of the entire dataset and the renamed Fuhrman variable are extracted into a file
4	The main research question was addressed using Classification Tree methodology. Using the 'rpart' package in the R language for statistical computing, a classification tree model was built using malignancy as the outcome variable and potential predictor variables as the covariates.[41] The covariates used were: age, symptoms at presentation, sex, degree of endophytic component, tumour volume, and central vs. peripheral location
5	The output is directed to specific files
6	The brief and detailed information derived from the tree are saved into a file
7	The confusion matrix generates measures of overall accuracy, sensitivity, specificity, positive and negative predictive values for the tree
8	The tree model was then validated using the 10-fold cross-validation routine in rpart and all the estimate of error rates are calculated for the 10 samples and then the mean estimate of the error rate is calculated
9	The tree was pruned using the 'tree' package in R in order to develop the most parsimonious tree that made clinical sense (simpler) without a significant change in the predictive ability (overall accuracy)
10	The confusion matrix was generated in order to obtain overall accuracy, sensitivity, specificity, positive and negative predictive values for the pruned tree
11	For the purposes of showing the results in a traditional framework a dummy variable with 5 levels was created. These levels correspond to the 5 terminal nodes of the final pruned tree. The variable was then used as a predictor in a logistic regression model with Benign as the outcome variable. Odds ratios with 95% confidence intervals were calculated using as the reference the terminal node in the tree that corresponded to: Tumour volume > 5.67 cc, Endophytic component > 35% and Incidental diagnosis

Appendix 3. R Code

```
library(rpart)
library(caret)
library(tree)

dat = indat

# outputs descriptive statistics to a file called desc.txt
sink("/Users/rrendon/Dropbox/THESIS/R FILES/desc.txt")
summary(dat)
summary(dat$Fuhrman)
sink()

# CART ANALYSIS

z.renalmass = rpart(benign ~ age + sex + incidental + Endophytic +
radVolume + peripheral, data=dat)
pdf("/Users/rrendon/Dropbox/THESIS/R FILES/z.renalmass.pdf")
plot(z.renalmass)
text(z.renalmass, all=TRUE, use.n=TRUE, cex=.6)
dev.off()

sink("/Users/rrendon/Dropbox/THESIS/R FILES/zrenalmass.txt")
z.renalmass
summary(z.renalmass)
# sensitivity and specificity
z.renalmass.pred = predict(z.renalmass, newdata=dat, type="class")
print("the cost of a FN is equal to the cost of a FP")
confusionMatrix(z.renalmass.pred, dat$benign, positive="Malignant")

# re-substitution estimate of error rate for the full tree
1-confusionMatrix(z.renalmass.pred, dat$benign,
positive="Malignant")[[3]][1]

# K-fold cross-validation estimate of error rate
# less biased (higher than re-sub estimate)
# (this code was obtained from
http://cours.zucker.fr/PROGRAMMES/rpartTitanic.R)
n = nrow(dat)
K = 10
foldsize = n%/%K
set.seed(5)
alea = runif(n)
rang = rank(alea)
bloc = (rang-1)%/%foldsize + 1
bloc = as.factor(bloc)
print(summary(bloc))

all.err = numeric(0)
for (k in 1:K){
    arbre = rpart(benign ~ age + sex + incidental + Endophytic +
radVolume + peripheral, data=dat[bloc!=k,], method="class")
    pred = predict(arbre, newdata=dat[bloc==k,],type="class")
    mc = table(dat$benign[bloc==k],pred)
    err = 1.0 - (mc[1,1]+mc[2,2])/sum(mc)
    all.err = rbind(all.err,err)
```

```
}
print(all.err)

err.cv = mean(all.err)
print(err.cv)
sink()

#PRUNED SUBTREE WITH 5 TERMINAL NODES
z.RenalMass <- tree(benign ~ age + sex + incidental + Endophytic +
radVolume + peripheral, data=dat,split="gini")
plot(z.RenalMass)
text(z.RenalMass)

snip.RenalMass <- snip.tree(z.RenalMass,nodes=c(4,5,6,14,15))
plot(snip.RenalMass)
text(snip.RenalMass)

sink("/Users/rrendon/Dropbox/THESIS/R FILES/SnipRenalMass.txt")
snip.RenalMass
summary(snip.RenalMass)
# sensitivity and specificity
snip.RenalMass.pred = predict(snip.RenalMass, type="class")
print("the cost of a FN is equal to the cost of a FP")
confusionMatrix(snip.RenalMass.pred, dat$benign, positive="Malignant")

# re-substitution estimate of error rate
1-confusionMatrix(snip.RenalMass.pred, dat$benign,
positive="Malignant")[[3]][1]
```

Appendix 4. SAS Code

```
proc format ;

value CARTvar     1 = "volLT5.67 EndoLT45"
                        2 = "volLT5.67 EndoGT45"
                        3 = "volGT5.67 EndoLT35"
                        4 = "volGT5.67 EndoGT35 Symptoms"
                        5 = "volGT5.67 EndoGT35 Incidental" ;

value Benign      0 = "Malignant"
                        1 = "Benign" ;

value incidental  0 = "Symptoms"
                                1 = "Incidental" ;

run;

data final ;
      infile '/Users/rrendon/Dropbox/THESIS/R FILES/ spreadsheet-110712
for SAS Final.csv' lrecl=2000 dlm=',' firstobs=2 truncover ;
      input ido idf age sex $ incidental hematuria pain radX radY radZ
            Upper Middle Lower Anterior Posterior Medial Lateral
            Endophytic CHP $ Benign PartialNeph Laparoscopic Stage $
            clearCell papillary chromophobe GranularCell angiomyolipoma
            oncocytoma benignCystic metanephricAdenoma leiomyoma
cysticNephroma
            SpindleCell vascInvasion fatInvasion Fuhrman Margins $
Size1 Size2 Size3
            histological Axis1 $ Axis2 $ Axis3 $ radVolume Peripheral $
;
      if radVolume < 5.66527 and Endophytic < 45 then CARTvar = 1 ;
      if radVolume < 5.66527 and Endophytic > 45 then CARTvar = 2 ;
      if radVolume > 5.66527 and Endophytic < 35 then CARTvar = 3 ;
      if radVolume > 5.66527 and Endophytic > 35 and incidental = 0
then CARTvar = 4 ;
      if radVolume > 5.66527 and Endophytic > 35 and incidental = 1
then CARTvar = 5 ;
      format CARTvar CARTvar. Benign Benign. incidental incidental. ;
run ;

proc logistic ;
      class CARTvar (ref="volGT5.67 EndoGT35 Incidental") ;
      model Benign (event="Benign") = CARTvar ;
run ;
```